DNA sequence, and recombinant preparation of the grass pollen allergen Lol p 4

Background of the invention

The present invention relates to the provision of a DNA sequence of the major grass pollen allergen Lol p 4. The invention also encompasses fragments, new combinations of partial sequences and point mutants having a hypoallergenic action. The recombinant DNA molecules and the derived polypeptides, fragments, new combinations of partial sequences and variants can be utilised for the therapy of pollen-allergic diseases. The proteins prepared by recombinant methods can be employed for *in vitro* and *in vivo* diagnosis of pollen allergies.

Type 1 allergies are of importance worldwide. Up to 20% of the population in industrialised countries suffer from complaints such as allergic rhinitis, conjunctivitis or bronchial asthma. These allergies are caused by allergens present in the air (aeroallergens) which are released by sources of various origin, such as plant pollen, mites, cats or dogs. Up to 40% of these type 1 allergy sufferers in turn exhibit specific IgE reactivity with grass pollen allergens (Freidhoff et al., 1986, J. Allergy Clin. Immunol. 78, 1190-2001).

The substances which trigger type 1 allergy are proteins, glycoproteins or polypeptides. After uptake via the mucous membranes, these allergens react with the IgE molecules bonded to the surface of mast cells in sensitised individuals. If two IgE molecules are crosslinked to one another by an allergen, this results in the release of mediators (for example histamine, prostaglandins) and cytokines by the effector cell and thus in the corresponding clinical symptoms.

A distinction is made between major and minor allergens, depending on the relative frequency with which the individual allergen molecules react with the IgE antibodies of allergy sufferers.

For perennial ryegrass (*Lolium perenne*), Lol p 1 has been identified as a major allergen (Freidhoff et al., 1986, J. Allergy Clin. 78:1190-1201) and its primary structure has been elucidated (Perez et al., 1990, J. Biol. Chem. 265:16210-16215). A further major allergen is Lol p 2 (Freidhoff et al., 1986, J. Allergy Clin. 78:1190-1201), the primary structure of which was described in 1993 (Ansari et al., 1989, J. Biol. Chem.: 264:11181-11185). A further major allergen of perennial ryegrass is Lol p 5 (Mattiesen and Löwenstein 1991, Clin. Exp. Allergy 21: 297-307). The primary structure of Lol p 5 is also known (Ong et al., 1993, Gene 134:235-240).
Perennial ryegrass furthermore contains the major allergens from groups 4 (Fahlbusch et al. 1998, Clin. Exp. Allergy 28: 799-807) and 13 (Petersen et al., 2001, J. Allergy Clin. Imm. 107:856-862).

Lol p 4 is a typical basic glycoprotein (Jaggi et al, 1989, Int. Arch. Allergy Appl. Immunol. 89:342-348, Jaggi et al., 1989, J. Allergy Clin. Immunol. 83:845-852) and is comparable with the well-studied Phl p 4, Cyn d 4 and Dac g 4 in terms of cross-reactivity with specific IgE antibodies (Haavik et al., 1985, Int. Arch. Allergy Appl. Immunol. 78:260-268; Su et al., 1991, Clin. Exp. Allergy 21:449-455; Leduc-Brodard et al., 1996, J. Allergy Clin. Immunol. 98:1065-1072; 14-17).

These homologous molecules from the *Poaceae* form allergen group 4, the molecules of which have high immunological cross-reactivity with one another both with monoclonal murine antibodies and also with human IgE antibodies (Fahlbusch et al., 1993 Clin. Exp. Allergy 23:51-60; Leduc-Brodard et al., 1996, J. Allergy Clin. Immunol. 98:1065-1072; Su et al., 1996, J. Allergy Clin. Immunol. 97:210; Fahlbusch et al., 1998, Clin. Exp.

Allergy 28:799-807; Gavrovic-Jankulovic et al., 2000, Invest. Allergol. Clin. Immunol. 10 (6):361-367; Stumvoll et al. 2002, Biol. Chem. 383:1383-1396; Grote et al., 2002, Biol. Chem. 383:1441-1445; Andersson and Lidholm, 2003, Int. Arch. Allergy Immunol. 130:87-107; Mari, 2003, Clin. Exp. Allergy, 33 (1):43-51).

In contrast to the above-mentioned major allergens Lol p 1, Lol p 2, Lol p 5 from *Lolium perenne*, the primary structure of Lol p 4 has not yet been elucidated.

From the group 4 allergen from *Dactylus glomerata* , it has hitherto only been possible for peptides to be obtained by enzymatic degradation and sequenced:
DIYNYMEPYVSK (SEQ ID NO 7),
VDPTDYFGNEQ (SEQ ID NO 8),
ARTAWVDSGAQLGELSY (SEQ ID NO 9)
and GVLFNIQYVNYWFAP (SEQ ID NO 10, Leduc-Brodard et al., 1996, J. Allergy Clin. Immunol. 98: 1065-1072).
Peptides have also been obtained from the group 4 allergen of sub-tropical Bermuda grass (*Cynodon dactylon*) by proteolysis and sequenced:
KTVKPLYIITP (SEQ ID NO 11),
KQVERDFLTSLTKDIPQLYLKS (SEQ ID NO 12),
TVKPLYIITPITAAMI (SEQ ID NO 13),
LRKYGTAADNVIDAKVVDAQGRLL (SEQ ID NO 14),
KWQTVAPALPDPNM (SEQ ID NO 15),
VTWIESVPYIPMGDK (SEQ ID NO 16),
GTVRDLLXRTSNIKAFGKY (SEQ ID NO 17),
TSNIKAFGKYKSDYVLEPIPKKS (SEQ ID NO 18),
YRDLDLGVNQVVG (SEQ ID NO 19),
SATPPTHRSGVLFNI (SEQ ID NO 20),

and AAAALPTQVTRDIYAFMTPYVSKNPRQAYVNYRDLD (SEQ ID NO 21, Liaw et al., 2001, Biochem. Biophys. Research Communication 280: 738-743).

For *Lolium perenne*, peptide fragments having the following sequences have been described for the basic group 4 allergen: FLEPVLGLIFPAGV (SEQ ID NO 22) and GLIEFPAGV (SEQ ID NO 23, Jaggi et al., 1989, Int. Arch. Allergy Appl. Immunol. 89: 342-348).

However, these peptide sequences which have been described for *Lolium perenne* and other group 4 allergens have hitherto not resulted in the elucidation of the primary structure of the Lol p 4 allergen.

As the first sequence of a group 4 allergen, the still unpublished sequence of Phl p 4 from *Phleum pratense* has been elucidated by the inventors of the present patent application and described in International Application WO 04/000881.

The object on which the present invention is based therefore consisted in the provision of a DNA sequence of the Lol p 4 gene encoding an allergen having the immunological properties of Lol p 4, and a corresponding recombinant DNA, on the basis of which the allergen can be expressed as protein and made available, as such or in modified form, for pharmacologically significant exploitation. The sequence of Phl p 4 was the starting point for the present invention.

List of sequences according to the invention

- DNA sequence from the Lol p 4 gene (SEQ ID NO 1).
- Protein sequence derived from the DNA sequence in accordance with SEQ ID NO 1 (SEQ ID NO 2).

- DNA sequence (SEQ ID NO 3), composed of nucleotides 1-200 of Phl p 4 (in accordance with SEQ ID NO 5), 201-1472 of Lol p 4 (in accordance with SEQ ID NO 1) and 1473-1503 of Phl p 4 (in accordance with SEQ ID NO 5).
- Protein sequence (SEQ ID NO 4), composed of amino acids 1-67 of Phl p 4 (in accordance with SEQ ID NO 6), 68-490 of Lol p 4 (in accordance with SEQ ID NO 2) and 491-500 of Phl p 4 (in accordance with SEQ ID NO 6) having the properties, in particular immunological properties, of Lol p 4, encoded by the DNA sequence in accordance with SEQ ID NO 3.
- DNA sequence of Phl p 4 (SEQ ID NO 5), in accordance with SEQ ID NO 5 from WO 04/000881.
- Protein sequence of Phl p 4 (SEQ ID NO 6), in accordance with SEQ ID NO 6 from WO 04/000881.

Description of the invention

The present invention provides for the first time a DNA sequence of the major grass pollen allergen Lol p 4 (SEQ ID NO 1) which encodes an allergen having the immunological properties of Lol p 4.
The present invention therefore relates to a DNA molecule encoding an allergen having the properties of Lol p 4, corresponding to a nucleotide sequence in accordance with SEQ ID NO 1.
The invention furthermore relates to a DNA molecule encoding an allergen having the properties of Lol p 4, corresponding to a nucleotide sequence in accordance with SEQ ID NO 3, composed of nucleotides 1-201 of Phl p 4 (in accordance with SEQ ID NO 5), 202-1470 of Lol p 4 (SEQ ID NO 1) and 1471-1500 of Phl p 4.

The invention furthermore relates to sequences homologous to the DNA sequences according to the invention and corresponding DNA molecules of

group 4 allergens from other *Poaceae*, such as, for example, *Dactylis glomerata, Poa pratensis, Cynodon dactylon, Holcus lanatus, Secale cerale, Triticum aestivum* and *Hordeum vulgare,* which, owing to the sequence homology that exists, hybridise with the DNA sequences according to the invention under stringent conditions, or have immunological cross-reactivity with respect to Lol p 4.

The invention also encompasses fragments, new combinations of partial sequences and point mutants having a hypoallergenic action.

The invention therefore furthermore relates to corresponding partial sequences, a combination of partial sequences, or replacement, elimination or addition mutants which encode an immunomodulatory, T-cell-reactive fragment of a group 4 allergen from the *Poaceae*.

With knowledge of the DNA sequence of the naturally occurring allergens, it is now possible to prepare these allergens as recombinant proteins which can be used in the diagnosis and therapy of allergic diseases (Scheiner and Kraft, 1995, Allergy 50: 384-391).

A classical approach to effective therapeutic treatment of allergies is specific immunotherapy or hyposensitisation (Fiebig, 1995, Allergo J. 4 (6): 336-339, Bousquet et al., 1998, J. Allergy Clin. Immunol. 102 (4): 558-562). In this method, the patient is injected subcutaneously with natural allergen extracts in increasing doses. However, there is a risk in this method of allergic reactions or even anaphylactic shock. In order to minimise these risks, innovative preparations in the form of allergoids are employed. These are chemically modified allergen extracts which have significantly reduced IgE reactivity, but identical T-cell reactivity compared with the untreated extract (Fiebig, 1995, Allergo J. 4 (7): 377-382).

Even more substantial therapy optimisation would be possible with allergens prepared by recombinant methods. Defined cocktails of high-purity allergens prepared by recombinant methods, optionally matched to the individual sensitisation patterns of the patients, could replace extracts from natural allergen sources since these, in addition to the various allergens, contain a relatively large number of immunogenic, but non-allergenic secondary proteins.
Realistic perspectives which may result in reliable hyposensitisation with expression products are offered by specifically mutated recombinant allergens in which IgE epitopes are specifically deleted without impairing the T-cell epitopes which are essential for therapy (Schramm et al., 1999, J. Immunol. 162: 2406-2414).

A further possibility for therapeutic influencing of the disturbed TH cell equilibrium in allergy sufferers is immunotherapeutic DNA vaccination, which involves treatment with expressible DNA which encodes the relevant allergens. Initial experimental evidence of allergen-specific influencing of the immune response has been furnished in rodents by injection of allergen-encoding DNA (Hsu et al., 1996, Nature Medicine 2 (5): 540-544).

The present invention therefore also relates to a DNA molecule described above or below as medicament and to a corresponding recombinant expression vector as medicament.

The corresponding proteins prepared by recombinant methods can be employed for therapy and for *in vitro* and *in vivo* diagnosis of pollen allergies.
For preparation of the recombinant allergen, the cloned nucleic acid is ligated into an expression vector, and this construct is expressed in a suitable host organism. After biochemical purification, this recombinant allergen is available for detection of IgE antibodies by established methods.

The present invention therefore furthermore relates to a recombinant expression vector comprising a DNA molecule described above or below, functionally linked to an expression control sequence, and a host organism transformed with said DNA molecule or said expression vector.

The invention also relates to the use of at least one DNA molecule described above or at least one expression vector described above for the preparation of a medicament for the immunotherapeutic DNA vaccination of patients with allergies in the triggering of which group 4 allergens from the *Poaceae,* in particular Lol p 4, are involved and/or for the prevention of such allergies.

As already stated, the invention can be used as an essential component in a recombinant allergen- or nucleic acid-containing preparation for specific immunotherapy. A number of possibilities arise here. On the one hand, the protein with an unchanged primary structure may be a constituent of the preparation. On the other hand, a hypoallergenic (allergoid) form can be used in accordance with the invention for therapy in order to avoid undesired side effects by specific deletion of IgE epitopes of the molecule as a whole or the production of individual fragments which encode T-cell epitopes. Finally, the nucleic acid per se, if ligated with a eukaryotic expression vector, gives a preparation which, when applied directly, modifies the allergic immune state in the therapeutic sense.

The present invention furthermore relates to polypeptides encoded by one or more of the DNA molecules described above, preferably in their property as medicament. In particular, the polypeptides are a protein corresponding to an amino acid sequence in accordance with SEQ ID NO 2 or a protein which contains this amino acid sequence or a part of this sequence, having the properties, in particular immunological properties, of Lol p 4, and a protein corresponding to an amino acid sequence in accordance with SEQ

ID NO 4 having the properties, in particular immunological properties, of Lol p 4.

Accordingly, the invention also relates to a process for the preparation of such polypeptides by cultivation of a host organism and isolation of the corresponding polypeptide from the culture.

The invention likewise relates to the use of at least one polypeptide or protein described above for the preparation of a medicament for the diagnosis and/or treatment of allergies in the triggering of which group 4 allergens from the *Poaceae*, in particular Lol p 4, are involved, and for the prevention of such allergies.

When determining the protein and DNA sequence of Lol p 4, the following procedure was followed:

The Lol p 4 DNA sequence in accordance with SEQ ID NO 1 according to the invention was amplified, cloned and sequenced by PCR with specific primers (Table 1) derived from the Phl p 4 sequence in accordance with SEQ ID NO 5 as described in WO 04/000881. A total of 7 clones were analysed. Analysis of the clones gave a uniform sequence. Three Lol p 4 DNA sequences were obtained by PCR with primers #87 and #83. The Lol p 4 DNA sequence amplified with these primers encodes the corresponding amino acids 68-401, based on the numbering of mature Phl p 4 in accordance with SEQ ID NO 6. Two further clones were obtained by PCR with primers #87 and #189. The Lol p 4 DNA sequence amplified with these primers encodes the corresponding amino acids 68-490 (numbering corresponding to Phl p 4 sequence). Two clones were obtained by PCR with primers #87 and #131. The amplified Lol p 4 DNA sequence likewise encodes the corresponding amino acids 68-490 (numbering corresponding to Phl p 4 sequence). Primers #131 and #189 correspond to the codons for the final 10 amino acids of the Phl p 4 protein and span the stop codon.

The DNA sequence in accordance with SEQ ID NO 3 according to the invention was obtained by methods known per se (PCR technique with overlapping primers).

For the preparation of the recombinant allergens according to the invention, the DNA sequences in accordance with SEQ ID NO 1 or 3 were incorporated into expression vectors (for example pProEx, pSE 380). For the N-terminal amino acids known from the protein sequencing, *E. coli*-optimised codons were used.

After transformation in *E. coli*, expression and purification of the recombinant allergens according to the invention by various separation techniques, the proteins obtained were subjected to a refolding process.
Both allergens can be employed for highly specific diagnosis of grass pollen allergies. This diagnosis can be carried out *in vitro* by detection of specific antibodies (IgE, IgG1 - 4, IgA) and reaction with IgE-loaded effector cells (for example basophiles from blood) or *in vivo* by skin test reactions and provocation at the reaction organ.

The reaction of the allergens according to the invention with T-lymphocytes from grass pollen allergy sufferers can be detected by allergen-specific stimulation of the T-lymphocytes for proliferation and cytokine synthesis both with T-cells in freshly prepared blood lymphocytes and also on established nLol p 4-reactive T-cell lines and clones.

The triplets encoding the cysteines were modified by site-specific mutagenesis in such a way that they encode other amino acids, preferably serine. Both variants in which individual cysteines have been replaced and those in which various combinations of 2 cysteine residues or all cysteines have been modified were prepared. The expressed proteins of these cysteine point mutants have greatly reduced or zero reactivity with IgE anti-

bodies from allergy sufferers, but react with the T-lymphocytes from these patients.

The present invention therefore furthermore relates to a DNA molecule described above or below in which one, a plurality of or all cysteine residues of the corresponding polypeptide have been replaced by another amino acid by site-specific mutagenesis.

The immunomodulatory activity of hypoallergenic fragments which correspond to polypeptides having T-cell epitopes and that of the hypoallergenic point mutants (for example cysteine replacements) can be detected by their reaction with T-cells from grass pollen allergy sufferers.

Such hypoallergenic fragments or point mutants of the cysteines can be employed as preparations for hyposensitisation of allergy sufferers since they react with the T-cells with equal effectiveness, but result in reduced IgE-mediated side effects owing to the reduced or entirely absent IgE reactivity.

If the nucleic acids encoding the hypoallergenic allergen variants according to the invention or the unmodified DNA molecules according to the invention are ligated with a human expression vector, these constructs can likewise be used as preparations for immunotherapy (DNA vaccination).

Finally, the present invention relates to pharmaceutical compositions comprising at least one DNA molecule described above or at least one expression vector described above and optionally further active ingredients and/or adjuvants for the immunotherapeutic DNA vaccination of patients with allergies in the triggering of which group 4 allergens from the *Poaceae,* in particular Lol p 4, are involved and/or for the prevention of such allergies.

A further group of pharmaceutical compositions according to the invention comprises at least one polypeptide described above instead of the DNA and is suitable for the diagnosis and/or treatment of said allergies.

Pharmaceutical compositions in the sense of the present invention comprise, as active ingredients, a polypeptide according to the invention or an expression vector and/or respective pharmaceutically usable derivatives thereof, including mixtures thereof in all ratios. The active ingredients according to the invention can be brought into a suitable dosage form here together with at least one solid, liquid and/or semi-liquid excipient or adjuvant and optionally in combination with one or more further active ingredients.
Particularly suitable adjuvants are immunostimulatory DNA or oligonucleotides having CpG motives.

These compositions can be used as therapeutic agents or diagnostic agents in human or veterinary medicine. Suitable excipients are organic or inorganic substances which are suitable for parenteral administration and do not adversely affect the action of the active ingredient according to the invention. Suitable for parenteral use are, in particular, solutions, preferably oil-based or aqueous solutions, furthermore suspensions, emulsions or implants. The active ingredient according to the invention may also be lyophilised and the resultant lyophilisates used, for example, for the preparation of injection preparations. The compositions indicated may be sterilised and/or comprise adjuvants, such as preservatives, stabilisers and/or wetting agents, emulsifiers, salts for modifying the osmotic pressure, buffer substances and/or a plurality of further active ingredients.
Furthermore, sustained-release preparations can be obtained by corresponding formulation of the active ingredient according to the invention – for example by adsorption on aluminium hydroxide.

The invention thus also serves for improving *in vitro* diagnosis as part of allergen component-triggering identification of the patient-specific sensitisation spectrum. The invention likewise serves for the preparation of significantly improved preparations for the specific immunotherapy of grass pollen allergies.

Table 1 Primers used

Primer number	SEQ ID NO	Sequence
#83	24	GGCTCCCGGGGCGAACCAGTAG
#87	25	ACCAACGCCTCCCACATCCAGTC
#131	26	GATAAGCTTGAATTCTGATTAGTACTTTTTGATCAGC GGCGGGATGCTC
#189	27	GATAAGCTTCTCGAGTGATTAGTACTTTTTGATCAGC GGCGGGATGCTC